Research Synopsis

On

Design and Analysis of Different Deep Learning Algorithms in Face Recognition

Submitted By
Bali Shankar Khurana

For the award of
Doctor of Philosophy in Computer Science

Under the guidance
of
DR. Ajay B. Kurhe
Assistant Professor (Computer Science)
Shri. Guru Buddhi Swami Mahavidyalaya, Purna District- Parbhani

Submitted to
SCHOOL OF COMPUTA TIONAL SCIENCES
SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY, NANDED

SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY, NANDED (M.S.)

2021-22

Table of Contents

1. **Abstract:** 3

2. **Introduction** 3

3. **Literature Review** 4

4. **Objectives of Research** 7

5. **Methodology** 7

6. **Significance of the Study** 9

7. **Proposed work Plan** 9

8. **References** 11

1. Abstract:

In the last few years face recognition has received a lot of attention. Recently face authentication has become very popular since it is easy to use and face recognition technologies have achieved good performance. Face biometric systems are widely used for recognizing the true identity of enrolled person on the basis of physiological appearances and behavioral patterns. However, this can be spoofed by nefarious users trying purposely to by-pass face recognition on system using various traits like photo or video of the person held in front of the camera may be accountable for performing face recognition. In this research work we mainly focus on optimizing face liveness detection approaches by using deep learning algorithms.

2. Introduction

The general public has immense need for security measures against spoof attack. Biometrics is the fastest growing segment of such security industry. Some of the familiar techniques for identification are facial recognition, fingerprint recognition, handwriting verification, hand geometry, retinal and iris scanner. Among these techniques, the one which has developed rapidly in recent years is face recognition technology and it is more direct, user friendly and convenient compared to other methods. Therefore, it has been applied to various security systems. But, in general, face recognition algorithms are not able to differentiate 'live' face from 'not live' face which is a major security issue. It is an easy way to spoof face recognition systems by facial pictures such as portrait photographs. Facial data used in biometrics can easily be stolen from social platforms or online communities and may be used for device spoofing. Playing video or placing a 3D dummy mask of the user in front of the camera may also be used to spoof the user's face. In order to guard against such spoofing, a secure system needs liveness detection. Integrated liveness detection may be used to solve certain problems where physiological signs of life are present in the database.

Face recognition is a kind of biometric identification technology based on the facial features of people. A series of technologies relate to face recognition by using cameras to collect images or video streams containing faces, and automatically detect and track faces in the images. Face verification is a subfield of face recognition, which refers to several images containing faces to judge whether these faces belong to the same identity. The method is to extract facial features from the target image to be verified, and then traverse the database of the known identity of facial features.

Due to the popularity of camera technology and the upsurge of machine learning, the basic research of face recognition has been relatively mature. According to a 2020 study by MarketsandMarkets, the global **biometric systems** market is expected to be worth USD 36.6 billion by 2021, and grow to USD 68.6 billion by 2025, with a CAGR of 13.4% during the forecast period (2021–2025).

The application of neural networks improves the performance of face recognition to a new level. The development trend is to build a deeper and more complex network to achieve higher accuracy, but the storage space and speed causes difficulties when attempting to meet the requirements of universal application. Face recognition includes multiple subtasks, which is often inefficient when the neural network-based face recognition algorithm is run locally on common devices. There are disadvantages such as high delay and poor user experience, and to date no good solution has been found. To solve this problem, a unified and efficient neural network algorithm for face recognition is required.

Liveness detection refers to determining whether biological information comes from a legitimate user of living organism when obtaining biological information. The method of liveness detection mainly distinguishes biometrics forged by nonliving substances, such as photographs, silica gel, and plastic, by identifying physiological information on the organism as a feature.

3. Literature Review

Face detection works to find the position of all faces in an image. Generally, the input is an image, and the output is the coordinates of any rectangular frames that detect the faces. Face detection is the basis of various face image analyses, and it is also the first step in the overall face verification algorithm. The main difficulties in face detection include: the diversity of facial gestures and angles, the influence of illumination intensity and angle, the possibility of partial occlusion of the face, the proximity of the face position affects the size of the face, and the face may appear in various positions of the image. At present, face detection algorithms can be divided into three categories: VJ framework-based, DPM model-based, and Convolutional Neural Network-Based.

In [1], the VJ framework was proposed by Viola and Jones in 2001. They used the integral graph to quickly calculate the Haar features of an image. The Haar features can reflect the contrast between parts of an image. The algorithm uses AdaBoost as a classifier and adopts a cascading

structure, which greatly improves the detection efficiency. Under the hardware conditions at that time, the processing speed can reach 15 fps. However, the VJ framework also has some shortcomings. Haar features are too simple and have insufficient stability. Moreover, the use of decision trees as weak classifiers can easily lead to overfitting. It also does not work well when the face is partially obscured or has an exaggerated expression

In [2], Deformable Part Model (DPM) is a variable component model. At the time of detection, the model first calculates the DPM feature map of the input image, and then the input image is unsampled by Gaussian pyramid to obtain an image twice as large as the original, and the DPM feature map is calculated. The model then uses the root filter to obtain a response map for the DPM feature map. The model uses the part filter for the feature map of the up sampled image, and then performs Gaussian pyramid down sampling on the obtained response graph. Thus, the response maps obtained by the root filter and the part filter has the same resolution. Finally, the response graph is weighted and averaged to obtain the final response graph. The DPM-based method can achieve better face detection than the VJ-based method in complex scenes such as outdoors, but the DPM model is still hard to realize real-time detection due to the computational complexity of the model.

In [3–8], Cascade CNN is a convolutional neural network implementation of the VJ framework. Cascade CNN uses three CNN cascading structures. Cascade CNN first constructs a detection image pyramid, uses the primary network scan to remove most of the windows, and then adjusts the window position and size through a correction network, and uses the non-maximum suppression to merge the height coincidence window as the next level network input for further detection. Cascade CNN solves the problem of illumination and angle better than traditional methods. However, the performance of the method is affected to some extent by the fact that the first-level network still uses the dense sliding window.

In face detection algorithms:

(1) The VJ framework improves the detection performance by cascading, but the face detection effect for complex expressions and angles is not good.

(2) The DPM model is less affected by noise, but is limited by the complex structure of the model. This leads to large calculations and poor real-time performance.

(3) Neural network-based detection methods can obtain better facial features, but these algorithms generally have the disadvantages of low interpretation, difficult parameter adjustment, and long operation time.

Liveness detection refers to determining whether biological information comes from a legitimate user of living organism when obtaining biological information. The method of liveness detection mainly distinguishes biometrics forged by nonliving substances, such as photographs, silica gel, and plastic, by identifying physiological information on the organism as a feature. The paper [9] divides the liveness detection technology into texture information analysis, motion information analysis, and living part analysis. The performance gap of the classifier based on the nonrigid motion of the authentic image, the noise difference, and the face background dependence are discussed. According to the input, liveness detection can be divided into single-frame input and continuous multiframe input.

The paper [10] proposes a liveness detection method that uses image distortion and color to construct feature vectors and uses SVM to perform two classifications. This method is not effective in the case where the forged facial image distortion is not serious. Color Texture **[11]** believes that the living and nonliving are indistinguishable in RGB space, but there are significant texture differences in other color spaces.

The paper **[12]** implements liveness detection based on Lambertian Reflection and believes that a true facial living body and a nonliving body from a video or a photo are differently reflected under the same lighting conditions.

The paper [13] enhances facial micromotion by inputting continuous multi-frame facial images, and then extracting dynamic texture features and histograms of oriented optical flow. Santosh Tirunagari et al. [14] used the dynamic mode decomposition DMD to obtain the subspace map of the maximum motion energy, then perform texture analysis, and finally, input to the SVM classifier for binary classification. This method has a poor effect on forging the video or shaking a printed image. The paper [15] introduces the liveness detection dataset PHOTO-ATTACK, which is extended based on PRINT-ATTACK, adding high-resolution screen images and mobile phone images.

In the past decade, Convolutional Neural Network (CNN) [16] has become one of the most popular techniques for solving computer vision problems. Many visual tasks, such as image classification, object detection, and face recognition, benefit from CNN's powerful learning and discriminative characterization. The CNN-based face recognition method usually regards CNN as a powerful feature extractor.

4. Objectives of Research

The objectives of the study provide the proper framework and layout of any research activities. Due to the paramountcy of face liveness detection the research will be carried out for the following objectives:

1. To study various issues in face recognition for biometric applications.
2. To address face liveness detection issue for face through texture analysis.
3. To develop/ modify existing method for improved Life Sign Detection.
4. Testing of developed technique/ method on existing database/ recorded images.
5. To evaluate the effectiveness of proposed deep learning model with other baseline models.

5. Methodology

In order to implement liveliness in face recognition, Researcher will use Python programming which has a rich library to build solutions much faster. For deep learning model design and training Researcher will use the TensorFlow framework and Keras which provide high level API binding to TensorFlow. Building and training models in Keras is much easier. For image reading we can use OpenCV library and NumPy library for array operations, like re-scaling image array, transforming array etc.

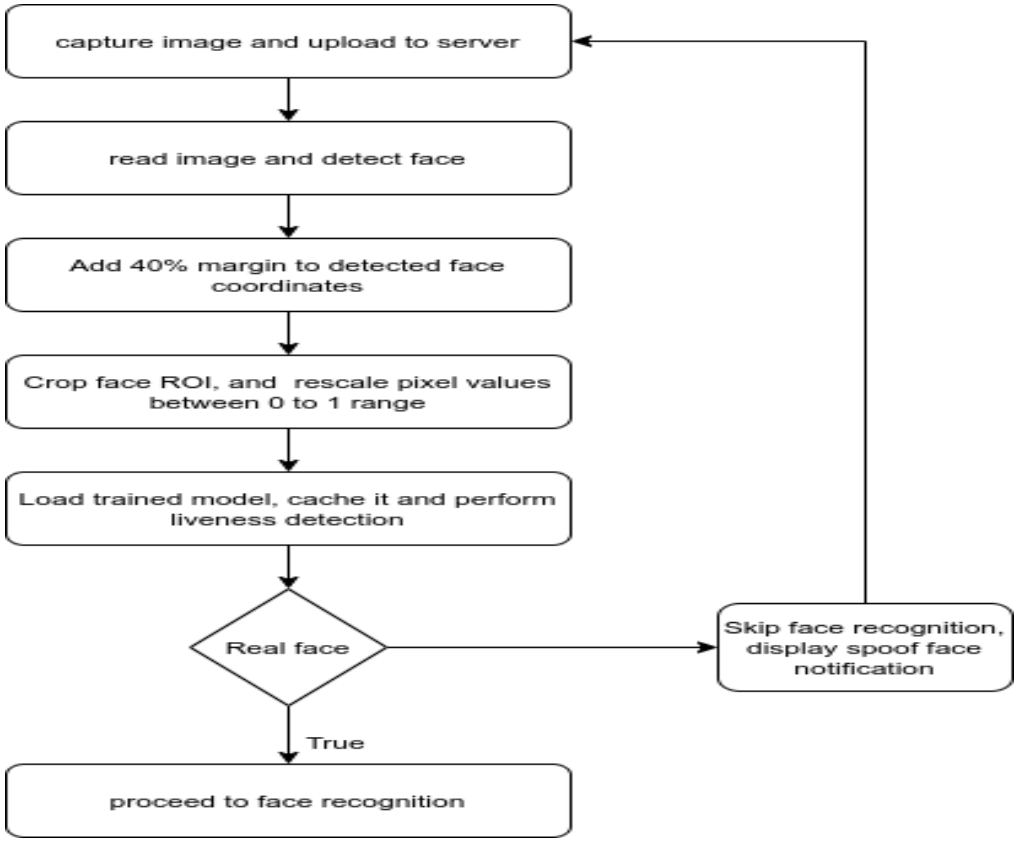
We will test CNN and ResNet architecture for liveness and find which give better results. Both of these architectures will require a lot of data to generalize the model. Model generalization will help us to reveal how a deep learning model behaves on unknown data, which necessitates a large amount of data for CNN/ResNet.

To generalize the model we will collect, process and train our framework with large number of images. The data will split into two parts: 80 percent for training and 20% for validation. To minimize training time, we will use transfer learning. For faster training and inference, a lightweight MobileNet model can also be used.

We have identified some large-scale dataset from public release. Raw image size: 1.2 TB from LCC-FASD extended, NUAA Imposter, CASIA anti spoof, CelebA-Spoof. CelebA-Spoof contains over 0.6 million images, and covers various light conditions, environments, and illumination conditions. CelebA-Spoof will be used to train and evaluate algorithms of face anti-spoofing. NUAA Photograph Imposter Database for face detection and geometric normalization. CASIA_faceAntispoof and LCC FASD datasets for false facial verification by using a photo, video. LCC-FASD extended dataset contains only live images captured using a mobile.

Work flow of the proposed framework

We have divided the approach into two parts, first face detection and second one liveness detection.


capture image and upload to server
read image and detect face
Add 40% margin to detected face coordinates
Crop face ROI, and rescale pixel values between 0 to 1 range
Load trained model, cache it and perform liveness detection
Real face
Skip face recognition, display spoof face notification
True
proceed to face recognition

First face detection will be performed on the user photo. Different face detectors such as Opencv Haar cascade, dlib's hog and CNN based detector, SSD-ResNet, Multi-task CNN and retina-face detector will be used. Retina face detects faces and returns a bounding box, after that the bounding box will be extended by a factor 0.4 to include the whole head, some background etc. based on extended bounding box coordinates face area will be cropped, resized to 300x300 and will be passed to CNN/ResNet model as input for liveness detection.

We plan to create our own dataset of 30 thousand images in size 300 x 300 to train our Neural Network to improve the performance of system and observe liveness detector was able to obtain overall 99% accuracy on validation set.

6. Significance of the Study

The traditional methods are mainly based on feature engineering. They rely on image quality evaluation, illumination, smoothness, and noise, and then obtain the test results through classifications. In addition to traditional detection methods, some literature has begun to apply deep learning to liveness detection in recent years but, limited by the small number of samples, the performance struggles to exceed traditional methods.

Deep learning approach to liveness detection introduced earlier simulates traditional liveness detection through CNN, but the effect is not good. Liveness detection can be divided into traditional methods based on image quality, texture information, and deep learning-based methods. The former is not effective in some special cases, while the latter is limited by the small number of data samples, which makes the network training difficult to fit, and the performance struggles to surpass the traditional method.

7. Proposed work Plan

In our work Researcher completely focus on finding new datasets which are publicly available on websites and study the factors which will determine face liveliness.

Tentative phase wise plan of proposed study we have designed as per the convenience of the study.

I YEAR	Completion of the coursework Review of literature Study various issues in face recognition for biometric applications Identification the problem Finalization of the research problem statement Formulation of the objectives, hypotheses and methodology
II YEAR	Tool preparation (Validity & Reliability of the algorithms) Pilot study To develop / modify existing method for improved Life Sign Detection. Obtain permission from the authorities and the subjects Data collection

	Analyzing the data and interpretation of study findings Summary preparation
III YEAR	Evaluate the effectiveness of proposed deep learning model with other baseline models. Submission of summary Conference proceedings Paper publications Paper presentations Writing the thesis Submission of thesis

8. References

1. *Viola, P.A.; Jones, M.J.; Snow, D. Detecting Pedestrians Using Patterns of Motion and Appearance. In Proceedings of the IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003.*
2. *Felzenszwalb, P.F.; Mcallester, D.A.; Ramanan, D. A discriminatively trained, multiscale, deformable part model. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.*
3. *Li, H.; Lin, Z.; Shen, X.; Brandt, J.; Hua, G. A convolutional neural network cascade for face detection. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.*
4. *Huang, L.; Yi, Y.; Deng, Y.; Yu, Y. DenseBox: Unifying Landmark Localization with End to End Object Detection. arXiv **2015**, arXiv:1509.04874.*
5. *Yang, S.; Luo, P.; Loy, C.-C.; Tang, X. From Facial Parts Responses to Face Detection: A Deep Learning Approach. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.*
6. *Jiang, H.; Learned-Miller, E. Face Detection with the Faster R-CNN. In Proceedings of the 2017 12th IEEE International Conference on Automatic Face and Gesture Recognition, Washington, DC, USA, 30 May–3 June 2017; pp. 650–657.*
7. *Chang, H.; Zhang, F.; Gao, G.; Zheng, H.; Chen, Y. Structure-constrained discriminative dictionary learning based on Schatten p-norm for face recognition. Digit. Signal Process. **2019**, 95, 102573. [CrossRef]*
8. *Zhou, Z.; He, Z.; Jia, Y.; Du, J.; Wang, L.; Chen, Z. Context prior-based with residual learning for face detection: A deep convolutional encoder-decoder network. Signal Process.-Image Commun. **2020**, 88, 115948. [CrossRef]*
9. *Kahm, O.; Damer, N. 2D face liveness detection: An overview. In Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 6–7 September 2012.*
10. *Di, W.; Hu, H.; Jain, A.K. Face Spoof Detection with Image Distortion Analysis. IEEE Trans. Inf. Forensics Secur. **2015**, 10, 746–761.*
11. *Boulkenafet, Z.; Komulainen, J.; Hadid, A. Face Spoofing Detection Using Color Texture Analysis. IEEE Trans. Inf. Forensics Secur. 2017, 11, 1818–1830. [CrossRef]*

12. *Tan, X.; Yi, L.; Liu, J.; Jiang, L. Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model. In European Conference on Computer Vision, Proceedings of the 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2020; Springer: Berlin/Heidelberg, Germany, 2020.*
13. *Bharadwaj, S.; Dhamecha, T.I.; Vatsa, M.; Singh, R. Face Anti-spoofing via Motion Magnification and Multifeature Videolet Aggregation. IEEE Trans. Inf. Forensics Secur. 2016, 3, 49–60.*
14. *Tirunagaris, S.; Poh, N.; Windridge, D.; Iorliam, A.; Suki, N.; Ho, A.T.S. Detection of Face Spoofing Using Visual Dynamics. IEEE Trans. Inf. Forensics Secur. 2015, 10, 762–777. [CrossRef]*
15. *Anjos, A.; Chakka, M.M.; Marcel, S. Motion-Based Counter-Measures to Photo Attacks in Face Recognition. IET Biom. 2014, 3, 147–158. [CrossRef]*
16. *Zamir, A.R.; Sax, A.; Shen, W.; Guibas, L.; Malik, J.; Savarese, S. Taskonomy: Disentangling Task Transfer Learning. In Proceedings of the IEEE/Cvf Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3712–3722.*
17. *Hsieh, Y.-H.; Leung, F.W. An overview of deep learning algorithms and water exchange in colonoscopy in improving adenoma detection. Expert Rev. Gastroenterol. Hepatol.* ***2019****, 13, 1153–1160. [CrossRef]*

Bali Shankar Khurana — **Research student**

Dr. Ajay B. Kurhe — **Research Guide**